

Subordination, subject-verb agreement and the CEFR

An analysis of linguistic complexity and accuracy in written L2 English and L3 French in relation to *the Common European Framework of Reference for Languages*

Author: Elin Nylander

English Studies

Autumn Term 2014

Course: ENGX64

Supervisor: Henrik Gyllstad

Centre for Languages and Literature

Lund University

Abstract

This investigation concerns written output produced by Swedish learners of English as a second language (L2) and French as a third language (L3). The analysed texts were rated according to the *Common European Framework of Reference for Languages* (CEFR), which is a widespread and language-independent document commonly used to assess communicative competence based on six main proficiency levels (A1, A2, B1, B2, C1 and C2). More specifically, the present study aims at comparing texts placed on different CEFR levels within the same language as well as samples from the same CEFR levels across different languages, in relation to linguistic complexity (i.e. the ability to use elaborate structures) and accuracy (i.e. correctness) (Housen & Kuiken, 2009). The targeted linguistic phenomena are subject-verb agreement and subordination, as it is hypothesised that the former is an indicator of accuracy, whereas the latter can be used to measure complexity (Norris & Ortega, 2009).

The data set consisted of email messages and narrative texts in English (N=24) and French (N=24) written by learners who were between 15 and 19 years old. All the data were manually coded for T-units (i.e. a measure referring to a main clause and all its associated dependent clauses) and clause-types by means of the CLAN toolbox (MacWhinney, 2000). This enabled a measuring of the complexity of the samples, as well as an examination of the kinds of subordinate clauses produced by the informants. The current investigation also contains an accuracy analysis based on the proportion of correct present tense verb forms in the texts.

According to the present findings, the analysed samples are relatively homogenous with regard to linguistic complexity, as the selected complexity measures only yielded statistically significant differences between the English B1 and the French A2 samples. However, the accuracy analysis showed that the percentage of correctly conjugated verbs in the texts generally increased with higher CEFR levels.

Table of Contents

1. Introduction	1
2. Background	2
2.1 The CEFR	2
2.2 The CAF framework in SLA research	4
2.3 The concept of interlanguage	5
2.4 The acquisition of L2 subordination and subject-verb agreement in English and French	
2.4.1 The acquisition of L2 subordination in English and French	7
2.4.2 The acquisition of subject-verb agreement in L2 English and French	10
2.5 Previous CEFR-related research	12
3. Methods and materials	16
3.1 Participants	16
3.2 The tasks	16
3.3 Coding and analysis of the data	18
4. Results	19
4.1 Differences in complexity between the CEFR levels	19
4.1.1 Complexity measures	19
4.1.2 Analysis of subordination types	22
4.2 Differences in accuracy between the CEFR levels	23
5. Discussion	24
6. Conclusion	28
References	30
Appendix 1: The rating scale used by the raters, compiled from several CEFR scales	33
Appendix 2: The instructions given to the informants	35

1. Introduction

The study reported in this essay aims to examine written texts rated according to the *Common European Framework of Reference for Languages* (CEFR). The CEFR is a document of reference in which, among other things, six levels of communicative competence (A1, A2, B1, B2, C1 and C2) are established and described. The framework was created in order to give European teachers, course designers and language learners a common language-independent basis for assessing language proficiency. Since it was first published in 2001, the CEFR has been translated into over 35 languages, and is now an internationally acknowledged document on which numerous language tests and curricula are based (Little, 2007).

However, although the framework has had enormous success, certain aspects of the CEFR have been questioned in studies investigating the document from a linguistic point of view. For example, Hulstjin (2007) states that the descriptions of linguistic competence included in the framework are based on observations made by language teachers rather than on actual data, and stresses, “the urgent need to empirically test the implications of the CEFR using real L2-learners” (p.8).

One investigation that contributes to the type of research called for by Hulstjin (2007) was carried out by Forsberg and Bartning (2010). Among other things, the study, which is based on written data in L2 French, aimed at determining whether morpho-syntactic features such as morphological errors in the verb phrase (VP) or the noun phrase (NP) can be used to discriminate between the CEFR levels. Forsberg and Bartning conclude that this is indeed possible and that, according to their data, “[m]easures of morpho-syntactic deviances [...] yield significant differences between the CEFR levels up to B2” (p. 151). In contrast, findings from a similar experiment conducted by Prodeau, Lopez and Véronique (2012) indicate that it is “difficult, if not impossible, to find lexical and grammatical means that [...] characterize only one [CEFR] level” (p.47).

Although the two studies mentioned above are different in that the latter has a more qualitative approach (see 2.5), the findings can be seen as inconclusive. With this and the need for more empirical underpinnings of the CEFR in mind, this essay contributes to the CEFR-related line of research by analysing written output in two different languages, and thus compares different CEFR levels within the same language as well as the same CEFR levels across different languages, as a means to relate the framework to empirical data. More

specifically, the aim of the present study is to investigate what learners of L2 English and L3 French at different CEFR levels produce in terms of subordination (i.e. the usage of dependent clauses) and subject-verb agreement, by comparing French and English texts which have all been given a CEFR rating. This type of comparison is called for because the language-independent approach in the CEFR seems to presuppose “that a particular communicative activity requires the same level of language proficiency in all languages involved” and “that communicative tasks and the proficiency needed to carry them out are learned in the same order and at approximately the same pace in all languages”, although this is not necessarily the case (Gyllstad, Granfeldt, Bernardini, & Källkvist, 2014, p.3). Moreover, the choice of linguistic phenomena (i.e. subordination and subject-verb agreement) is motivated by the fact that they are considered to reflect learners’ linguistic complexity (i.e. the ability to use elaborate structures) and accuracy (i.e. correctness). Together with fluency (i.e. eloquence), these notions form the so-called CAF framework, which is commonly used to measure linguistic competence, and where *C* stands for *complexity*, *A* for *accuracy* and *F* for *fluency* (Housen & Kuiken, 2009). However, according to Gunnarsson (2012), most of the current research related to the CAF concerns oral output. This calls for studies examining written data, in order to expand the application of the CAF framework, a need that the current investigation addresses.

In the following section, a background to the present study is outlined. It contains a more detailed presentation of the CAF framework mentioned above, and describes essential features of the acquisition of L2 subordination and subject-verb agreement in English and French. Section 2 also includes a summary of previous studies related to the CEFR scale, and leads up to the research questions addressed in this essay, which concern the differences between texts from the targeted CEFR levels in terms of linguistic complexity and accuracy. Firstly, however, we will take a look at some of the criteria included in the CEFR.

2. Background

2.1 The CEFR

The CEFR is an exhaustive document containing language-independent descriptive scales used to assess language skills. It includes criteria presented according to six main proficiency levels (A1-C2), which, in turn, align with three proficiency bands: Basic user (A),

Independent user (B) and Advanced user (C). The framework has an “action-oriented” approach, and language learners are regarded as “social agents” who “have tasks (not exclusively language-related) to accomplish in a given set of circumstances, in a specific environment, and within a particular field of action” (Council of Europe, 2001, p. 9). This is for instance reflected in the fourth chapter of the framework, which includes functional scales focusing on what learners at the different levels can do and how well they can do it. The scales are related to different domains of language use (e.g. personal or public) and language activities (e.g. written interaction). Further, the fifth chapter of the CEFR outlines the linguistic competences needed to carry out the different language activities referred to in chapter 4, and includes criteria specifying what learners at the respective levels must be able to master in terms of, for instance, grammatical accuracy and phonological control (see Appendix 1) (Council of Europe, 2001).

We saw earlier that despite playing an important role in language education, there are some theoretical issues related to the CEFR. For example, Gyllstad et al. (2014) mention the fact that the descriptions of linguistic competence included in the framework are somewhat vague and not clearly related to current second language acquisition (SLA) research. They also point out that although not explicitly stated, an underlying assumption in the CEFR is that learners who are considered being at, for instance, level B1, perform at this level with respect to both the functional scales in chapter 4 and the competence-based criteria in chapter 5. However, according to Hulstijn (2007), this is not always the case, as some learners know how to perform few language tasks very well, while others can do a wide range of things but with less quality, and only some language users are as proficient from a qualitative perspective as they are in terms of linguistic quantity.

The terminology used to describe communicative proficiency in the CEFR is also noteworthy, as certain descriptors frequently re-occur. For example, the framework commonly refers to the notion of complexity, and talks about *complex language* and *complex speech* as well as the opposite (e.g. *simple structures*). When analysing the Swedish translation of the CEFR, Gyllstad and colleagues (2014) found 59 occurrences of the adjective *complex* and 90 examples of *simple* or *basic*. They also observed a difference in the distribution of these adjectives, and report that the descriptions of proficiency below the level B1 include the adjectives *simple* and *basic* whereas the words *simple* and *complex* are used at the B1 level and only *complex* is employed at the levels B2, C1 and C2. According to Gyllstad et al. (2014), “it is thus possible to interpret level B1 as the threshold level, where the CEFR assumes that the users recognise their linguistic system in such a way that it

becomes increasingly complex” (p. 7).

Further, words like *accurate* and *accuracy* are also used to define language proficiency in the CEFR. For instance, learners placed at the CEFR level B1 are expected to be able to “[c]ommunicate *with reasonable accuracy* in familiar contexts [...]” (Council of Europe, 2001, p. 114, my emphasis). The problem, however, is that, to my knowledge, the framework provides a clear definition of what neither accuracy nor complexity actually is. This is worth stressing because as we will see in sub-section 2.2, linguistic accuracy and complexity are central constructs within the field of SLA.

2.2 The CAF framework in SLA research

As stated in the introduction, L2 proficiency can be described and measured in terms of complexity, accuracy and fluency (i.e. the CAF triad). A common definition of complexity is “the ability to use a wide and varied range of sophisticated structures and vocabulary in the L2” whereas accuracy is about “produc[ing] target-like and error-free language”. Finally, fluency can be defined as the ability to “produce the L2 with native-like rapidity, pausing, hesitation or reformulation” (Housen, Kuiken, & Vedder, 2012, p. 2). The three components of the CAF framework are a common point of departure in SLA research (see e.g. Gunnarsson, 2012; Ågren, Granfeldt, & Schlyter, 2012; Kuiken & Vedder, 2012). They can, for instance, be used as dependent research variables, in which case they can be examined in relation to factors like age or learning contexts. Lately, however, researchers have also shown an interest in using complexity, accuracy and fluency as separate variables, and thus seek to explore their role in L2 acquisition in general.

However, as pointed out by for instance Norris and Ortega (2009), the CAF framework is currently facing several challenges. For example, Housen and Kuiken (2009) report that although the working definitions of complexity, accuracy and fluency provided above have been used since the 1990s, it is not unproblematic to pinpoint exactly what these notions entail, as different definitions sometimes are used for the same component of the framework. Starting with accuracy, this dimension of L2 proficiency is considered relatively straightforward, as it is generally seen as referring to correctness and “the extent to which an L2 learners’ performance [...] deviates from the norm (i.e. usually the native speaker). Such deviations from the norm are traditionally labelled ‘errors’” (Housen et al., 2012, p.4). However, this definition can also be considered ambiguous, as it raises the question of which languages norms learners should aim for, since some utterances may be considered erroneous

in certain communities and acceptable in others. Further, the construct *fluency* is typically used to talk about general eloquence, but is sometimes employed in a narrower manner by referring to, for instance, the number of repetitions uttered by a learner, which is why it is not always obvious what it means. Finally, complexity is typically the most difficult construct to define, as it can be used in many different contexts. For example, Housen and Kuiken (2009) point out the importance of differentiating between task complexity, which concerns the features of a language task, and L2 complexity, which is about the features of a learner's language proficiency. L2 complexity can, in turn, be divided into two sub-categories: linguistic complexity (i.e. the elaborateness of the linguistic system) and cognitive complexity, (i.e. the difficulty with which learners acquire different L2 features). Considering that the purpose of this study is to, among other things, examine the morpho-syntactic characteristics of written output, the present investigation addresses linguistic complexity.

Housen and Kuiken (2009) also stress that although the components of the CAF triad can be seen as separate proficiency dimensions, attention must also be paid to how they interact. The authors report that according to some research, learners sometimes develop linguistic accuracy at the cost of complexity, as it might be too demanding to produce output that is both complex and accurate.

Another point of discussion related to the CAF concerns the way the three components can be measured and assessed. According to Norris and Ortega (2009), focusing on subject-verb agreement is a way to measure accuracy, whereas the number of subordinate clauses in a text can function as an indicator of linguistic complexity. In this essay, it is assumed that these features play an important role in the language learning process, as they are of decisive importance for learners' ability to communicate in the L2 (see e.g. Ågren, 2008; Bartning & Schlyter, 2004). For these reasons, subject-verb agreement and subordination are the two targeted language features in this essay, and later they will be examined in relation to the CEFR. Now, however, we turn to the construct *interlanguage*.

2.3 The concept of interlanguage

The term *interlanguage* (IL) was initially coined by the American linguist Larry Selinker (1972), and is used to refer to “the language system that each learner constructs at any given point in development [...]” (Ortega, 2009, p. 141). In other words, an IL can be seen as a

learner-specific variety of a target language (TL), and is thus a separate linguistic system that differs from both the TL and the native language of a learner (Tarone, 2006).

A central aspect of an IL is that of *fossilisation*, i.e. “the process in which the learner’s interlanguage stops developing, apparently permanently” (Tarone, 2006, p. 748). Sometimes, an IL can fossilise completely, in which case the learner in question is unable to move beyond a certain level of proficiency in the TL, despite being both motivated to learn and exposed to continuous TL input. In other cases, only one area of the IL fossilises, and a learner might, for example, be unable to master article usage, but might on the other hand demonstrate an advanced proficiency level in all other respects (Ortega, 2009). According to Tarone (2006), the concept of IL is typically believed to be characteristic for adults starting to learn a second language after puberty, as children acquiring an L2 are expected to be able to avoid fossilisation. Similarly, “[s]econd-language learners who begin their study of the [L2] after puberty [typically] do not succeed in developing a linguistic system that approaches that developed by children acquiring that language natively” (Tarone, 2006, p. 748).

Moreover, Selinker (1972) identified various psycholinguistic processes that are expected to shape learners’ IL. According to Ortega (2009), one such process is *simplification*, which entails a “one-meaning-one-form mapping” in that learners only use one of the existing forms of certain elements of a language” (p. 116). The IL can also be characterized by *overgeneralisation*. When this happens, the language learner uses a rule that he or she has learnt (e.g. that English verbs in the past tense end in *-ed*) in too many contexts, including those where the rule does not apply, by saying things like **goed* instead of *went*. Further, *restructuring* is a process concerning “self-reorganisation of grammar representations”, which means that learners make qualitative changes to the knowledge that they already possess (Ortega, 2009, p. 177). Those developing an IL can also evince so-called *u-shaped behaviour*, which is when advanced and native-like constructions are visible early in the acquisition process and then disappear, only to be used again at a later stage of development (Ortega, 2009).

Processes like simplification and u-shaped behaviour can be manifested in both oral and written L2 output in the form of various non-native-like utterances. However, it should be stressed that these features mirror the development of the IL, and should thus not always be looked upon negatively. For example, Ortega (2009) notes that the sentence **I didn’t went to Costa Rica* exemplifies a “better [...] [and] more advanced ‘error’” than **I will don’t see you tomorrow*, although both utterances are typically considered erroneous (p. 121). The next sub-

section accounts for the development of an IL with regard to subject-verb agreement and subordination.

2.4 The acquisition of L2 subordination and subject-verb agreement in English and French ¹

2.4.1 The acquisition of L2 subordination in English and French

Subordination can be described as a relationship between a main clause and a subordinate clause (i.e. dependent clause). Subordinate clauses are part of the main clause, and they can, for example, function as subjects or objects (Riegel, Pellat & Rioul, 1994). There are several main types of subordinate clauses, which will be outlined in this section based on the information provided in the *Longman Grammar of Spoken and Written English*, *Bonniers Franska Grammatik* and *Grammaire méthodique du français* (see References).

The first main type of subordinate clauses is adverbial clauses, which typically function as adverbials in a main clause. Thus, adverbial clauses can, for instance, add circumstantial information to the main clause they occur in, and they are often introduced by subordinators like *if* or *when*. The examples in (1)-(4) below all contain adverbial clauses, which have been underlined. Example (1) is a temporal adverbial clause (i.e. a circumstance adverbial related to time) and (2) is a causal adverbial clause expressing a reason. Moreover, example (3) is a concessive clause, which thus conveys a concession, and (4) is a conditional clause. Other types of adverbial clauses include clauses of purpose (e.g. *I save money so that I can travel during the holidays*) and clauses of result (e.g. *He was a student so he did not have a lot of money*) (Biber, Johansson, Leech, Conrad, & Finegan, 1999). Since the present study concerns subordination in L2 English and French, each English example is followed by a French translation. The sentences in (1)-(6) and the other examples mentioned in this section (2.4) are my own examples based on Biber et al. (1999) and Riegel et al. (1994), except for (7), which is taken from Wall, Ekman, Béhar, and Kroning (1999).

(1) I listen to music when I cook.

(a) J'écoute de la musique quand je fais la cuisine.

(b) J'écoute de la musique en faisant la cuisine.

¹ In this section, the abbreviation L2 does not necessarily refer to a learner's second language, but is also used when talking about the acquisition of other foreign languages (cf. Norris & Ortega, 2009).

(2) He wanted to have lunch because he was hungry.

Il voulait déjeuner parce qu'il avait faim.

(3) Although she was very intelligent, she did not have good grades.

Bien qu'elle soit très intelligente, elle n'avait pas de bonnes notes.

(4) If I had a lot of money, I would buy car.

Si j'avais beaucoup d'argent, j'achèterais une voiture

Moreover, it should be noted that the English temporal adverbial clause in (1) has two French equivalents. This is because the English example can be translated either with a subordinate clause introduced by *quand* (*when*) as in (1a) or with a construction called *gérondif* (e.g. *en faisant la cuisine*) as illustrated in (1b).

Further, nominal clauses can function as a subject, subject predicative or direct object in a main clause. This is exemplified in (5) and (6), where the underlined nominal clauses both function as a direct object. The sentences in (5) are examples of complement clauses (i.e. that-clauses) in English and French, and (6) exemplifies dependent interrogative clauses. Dependent interrogative clauses are introduced by a *wh*-word and are typically combined with verbs like *ask* and *wonder* to convey indirect questions. However, they can, for instance, also have an exclamative function as in *Now I know what a great job I have*. Examples of other types of nominal clauses are non-finite clauses with a verb in the *ing*-form (e.g. *Being a teacher is great*) and non-finite clauses with a *to*-infinitive (e.g. *It is unusual for a young lady to treat people this way*) (Biber et al., 1999).

(5) I think that I love you.

Je pense que je t'aime.

(6) I wonder where Pierre is going.

Je me demande où va Pierre.

Relative clauses constitute the third type of subordinate clauses referred to in this essay. They typically function as postmodifiers in a noun phrase, and are often introduced by a relative pronoun (i.e. a relativizer), which can also be omitted. In contrast to complement clauses such as the example in (5), relative clauses are “incomplete” as they have a “gap” in the form of a missing constituent corresponding to the meaning of the head noun in the main clause. For example, the (underlined) relative clause in *Pierre fetched the book that was in the kitchen* has

a gap in subject position, since the “underlying meaning” of the clause is *the book was in the kitchen*, even though *the book* “is not mentioned in the clause itself” (Biber et al., 1999, p.645).

In French, the main simple relative pronouns introducing relative clauses are *qui*, *que*, *quoi*, *dont* and *où*. There are also other relativizers such as *duquel*, which can be used with certain verbs (Riegel et al., 1994). The relative pronoun that is used depends on the function of the pronoun in the clause. For example, *qui* is typically used if the relative pronoun functions as a subject, whereas *que* is employed when the pronoun functions as an object (Wall et al., 1999). Thus, one relative pronoun in English (e.g. *that*) does not always correspond to the same relative pronoun in French (e.g. *que*). For instance, the French translation equivalent of *It is the book that I prefer* is *C’est le livre que je préfère*, whereas the English sentence *It is the book that I need* translates as *C’est le livre dont j’ai besoin*. Together, these examples show that English relative pronoun *that* can, for example, be translated as either *que* or *dont* depending on the context. Similarly, example (7) below illustrates how the French relativizer *qui* can correspond to two different relative pronouns in English (*that* and *who*) (Wall et al., 1999).

(7) “S’il te plaît, passe-moi les journeaux qui sont sur la table” (Wall et al., 1999, p.105).

Please pass me the papers that are on the table (my translation).

”Marie a un frère qui est vétérinaire “(Wall et al., 1999, p.104).

Marie has a brother who is a veterinarian (my translation).

Further, some clauses can be considered to be neither adverbial nor nominal or relative clauses. Comparative clauses (e.g. *She sings better than she dances*) and comment clauses (e.g. *It was a good idea I think*) are examples of subordinate clauses classified as *other* in the present study.

As to the acquisition of subordination in L2 English, Hawkins and Filipović (2012) summarize the findings obtained by researchers belonging to the *English Profile research Programme* who, based on extensive performance data from the *Cambridge Learner Corpus* aim at determining which English language features (e.g. types of dependent clauses) are characteristic for the respective CEFR levels. In terms of relevance to the present study, the findings presented by Hawkins and Filipović (2012) are somewhat limited, although they do report that finite complement clauses like the example in (5) are attested at the level A2.

They also state that usage of dependent interrogative clauses (see example 6) is one of several characteristics of B1 texts, whereas ing-clauses preceding the main clause as in *Talking about spare time, I think we should go to the zoo* are attested at the B2 level.

Moreover, Bartning and Schlyter (2004) investigated the acquisition of several grammatical phenomena in oral L2 French, including negation, subject-verb agreement and subordination. Based on previous findings regarding the respective language features, the authors suggest six developmental stages ranging from an initial stage (*le stade initial*) to a near-native stage (*le stade avancé supérieur*) that learners hypothetically go through when acquiring French as a second language. In terms of subordination, Bartning and Schlyter (2004) suggest that learners initially produce simple co-ordinated utterances containing connectors like *mais (but)* or *et (and)*. They then gradually acquire temporal, causal and relative clauses introduced by subordinators like *quand (when)* and *parce que (because)*. At the next stage, learners start to produce complement clauses, which they eventually learn to combine with other types of subordinate clauses. Later, language learners also use language features such as the *gérondif* mentioned above. Then, at the final stage, the utterances containing subordination are, among other things, characterized by a native-like employment of the subordinator *donc (so)*. We will return to the Bartning and Schlyter model in the next sub-section, which concerns the acquisition of subject-verb agreement in English and French.

2.4.2 The acquisition of subject-verb agreement in L2 English and French

In English, the notion of subject-verb agreement (i.e. subject-verb concord) is used to refer to the principle that “[a] subject and [a] verb phrase agree in number and person”, which means that the form of a verb that is used depends on which subject it refers to (Biber et al., 1999 p. 180). In English, the principal rule regarding this language feature is that when occurring with a third person singular subject (e.g. *he*), the *s*-form of lexical verbs (e.g. *talks*) and primary auxiliaries (e.g. *is* or *has*) is used. However, there is no subject-verb agreement with modal auxiliaries (e.g. **he cans speak English*), verbs in non-finite clauses (see 2.4.1), verbs in the imperative (e.g. *sit!*) or with subjunctive forms (e.g. *were*) (Biber et al., 1999).

Källkvist and Petersson (2006) looked at 17- and 14-year-old Swedish learners of L2 English and their ability to formulate a rule for subject-verb agreement. The authors notably found that 54 % of the 17-year-old informants participating in the study were “unable to verbalise an acceptable and simplified rule for the use of [the verb forms] *get* and *gets*”, despite having been exposed to several years of classroom instruction in English (p. 131). For

example 23 % of the 17-year-olds linked the verb form *gets* to subjects in the plural (e.g. **they gets*). Källkvist and Petersson (2006) suggest that this mistake might be related to the fact that the third person singular *-s* both sounds and is spelt the same way as the plural *-s* (e.g. *many cats*), as this might make learners associate the suffix *-s* with plurality, although the English *-s* can function as both a third person singular marker, a plural marker on nouns, and a marker of the possessive.

In French, the third person singular verb form in the present tense is not salient with regard to the present tense verb forms used with the other pronouns (or lexical noun phrases) like in English. Instead, French verbs tend to have different verb forms corresponding to different personal pronouns (e.g. *I, you, he* etc.) (Riegel et al., 1994). The complexity of French VP morphology is illustrated in Table 1 below, which summarizes how the French verbs *être* (to be), *parler* (to speak), *prendre* (to take) and *faire* (to do) are conjugated in the present tense compared to the corresponding English verb forms.

Table 1: The present tense verb forms of French verbs *être*, *parler*, *prendre* and *faire* and their English equivalents

A. Être (to be)		B. Parler (to speak) (regular)	
Être	To be	Parler	To speak
Je suis	I am	Je parle	I speak
Tu es	You are	Tu parles	You speak
Il/elle/on est	He/she/it is	Il elle/on parle	He/she/it speaks
Nous sommes	We are	Nous parlons	We speak
Vous êtes	You are	Vous parlez	You speak
Ils/elles sont	They are	Ils/elles parlent	They speak
C. Prendre (to take) (irregular)		D. Faire (to do)	
Prendre	To take	Faire	To do
Je prends	I take	Je fais	I do
Tu prends	You take	Tu fais	You do
Il/elle/on prend	He/she/it takes	Il/elle/on fait	He/she/it does
Nous prenons	We take	Nous faisons	We do
Vous prenez	You take	Vous faites	You do
Ils/elles prennent	They take	Ils elles font	They do

(Brossier, 2006)

In the previously mentioned study by Bartning and Schlyter (2004), the authors report that according to their framework, learners at the initial stage tend to master subject-verb agreement of common auxiliary verbs such as *être* (*to be*) in the first and third person singular (see Table 1). They then start to conjugate verbs with the first person plural ending with *-ons* such as *nous parlons* (*we speak*). At the next stage, learners typically start using verbs in the third person plural ending in *-ont* (e.g. *ils sont*), and as they become more proficient, they start producing irregular verbs like *ils prennent* (*they take*). Near-native speakers at the most advanced stage typically know how to conjugate irregular verbs correctly. Moreover, Bartning and Schlyter (2004) point out that the usage of constructions like *ils prennent* might be affected by the fact that the third person plural ending of the verb (*-ent*) is inaudible in oral French, which might result in learners overusing the non-marked form of the verb (**ils prend*). The role of audibility in the acquisition of French verb forms is also pointed out by Ågren (2008), who suggests that the fact that some French verb endings are pronounced in spoken language while others are not might cause problems for the development of learners' written IL as well, as this means that they must learn to distinguish between oral and written VP morphology. In a study on the acquisition of number agreement in written L2 French, Ågren (2008) used the stages proposed by Bartning and Schlyter (2004) as a point of departure. The author found that although the Bartning and Schlyter framework is based on oral French, it is also applicable to written data, as significant differences were observed when comparing output written by learners from the different stages. Moreover, returning to the CEFR scale introduced in section 1, this framework has been analysed in relation to the Bartning and Schlyter model in an investigation carried out by Rosenberg (2011), which will be summarized in the next section.

2.5 Previous CEFR-related research

As mentioned earlier, there is a need for more research related to the CEFR scale, although some studies have already contributed to the empirical underpinning of the framework (see e.g. Martin, Mustonen, Reiman, & Seilonen, 2010; Rosenberg, 2011; Gyllstad et al., 2014; Forsberg & Bartning, 2010; Prodeau et al., 2012; Hulstijn, Schoonen, de Jong, Steinel, & Florijn, 2011) Out of these, four investigations are particularly relevant for the present study, and will therefore be summarized in this sub-section.

Rosenberg (2011) examined the relationship between the developmental stages

proposed by Bartning and Schlyter (2004) (see 2.4.1) and the six main levels of the CEFR. The data consisted of 40 French texts written by 10 fifteen-year-old and 10 eighteen-year-old Swedish learners of L3 French, who had all produced one e-mail and one narrative text each. The samples came from a data set collected in connection with a study by Gyllstad et al. (2014) summarized below, and some of these texts will also be included in the present analysis (see 3.2). Rosenberg (2011) found that in general, the CEFR levels corresponded to Bartning and Schlyter's (2004) stages in that students whose productions contained features related to the initial stages (1 and 2) of Bartning and Schlyter's framework were typically given a CEFR rating between A1 and A2, whereas those who were on the intermediate stage (3) were placed between A2 and B1 on the CEFR scale. Finally, the learners whose grammatical profiles corresponded to the fourth stage (*avancé bas*) of the Bartning and Schlyter model were typically given a CEFR rating ranging from A2 to C1. Rosenberg's (2011) investigation is a welcome contribution to the empirical CEFR-related research requested by Hulstijn (2007) (see section 1), and it would be interesting to conduct a similar study with a different approach. Thus, the present study differs from Rosenberg's (2011) analysis, as it does not primarily focus on the developmental stages suggested by Bartning and Schlyter (2004). Instead, the CAF framework summarized in 2.2 is used to measure the informant's language proficiency.

Forsberg and Bartning (2010) examined the linguistic characteristics of the CEFR levels in terms of morpho-syntax, discursive features (e.g. relative pronouns), and the use of formulaic sequences (e.g. idiomatic expressions). More precisely, they investigated whether these language features could be used to discriminate between texts from the different levels of the CEFR. Forty-two Swedish university-students learning French as a foreign language participated in the study, and were placed on the CEFR scale by means of a vocabulary placement test, a self-assessment task and a written proficiency test. The students were then asked to (a) write a summary of a film or a book and (b) complete a second written task (e.g. an argumentative text) that was specific for each CEFR level. A professional CEFR rater also examined the texts in order to verify that the productions corresponded to the CEFR levels that the students had been placed on. In terms of relevance to the present study, the findings indicate significant morpho-syntactic differences between texts from the different CEFR levels up to B2, as errors regarding VP and NP morphology tended to “decrease with higher levels of CEFR, indicating an increase in students' accuracy” (Forsberg & Bartning, 2010, p. 142). The investigation carried out by Forsberg and Bartning offers valuable insight into the relationship between linguistic accuracy and the levels of CEFR scale. However, similarly to

Rosenberg's (2011) analysis, Forsberg and Bartning's (2010) study only looks at one language (French), which raises the question of whether the observed increase in accuracy can be found in texts written in other languages as well. This investigation therefore examines both English and French data as a means to compare texts from the same CEFR levels in the two different languages.

In a more recent study, Gyllstad et al. (2014) analysed written output in L2 English (N= 54), L3 French (N=38) and L4 Italian (N=28) produced by Swedish informants. The learners of English and French were between 10 and 19 years old and the learners of Italian were between 18 and 67 years of age. Moreover, the samples, which were in the form of narrative texts and email messages, had all been assigned to one of the CEFR levels by experienced raters. This enabled an investigation of the relationship between the rated CEFR levels and three different measures of syntactic complexity, namely the number of words per T-unit (i.e. a measure referring to a main clause and all its associated dependent clauses), the number of subordinate clauses per T-unit and the number of words per clause in each text. An additional aim was to determine the differences between (a) texts written in different languages and placed on the same CEFR level, and (b) texts placed on different CEFR levels within the same language, with respect to syntactic complexity. Gyllstad et al. (2014) observed "significant correlations [...] for all complexity measures across all three languages" (p. 16), which means that texts from the higher CEFR levels were generally more complex than samples from the lower levels. The findings also indicate a difference in syntactic complexity across languages in the B1 and B2 texts, whereas the texts from the levels A1 and A2 were "rather homogeneous" (p.22). As to the differences between CEFR levels within the same language, the authors report that it was possible to discriminate between A and B levels in English and French using measurements of mean length of T-unit and mean number of subordinate clauses per T-unit, whereas measurements of mean length of clause only could be used to differentiate between English levels. The study by Gyllstad and colleagues (2014) is a valuable underpinning of the CEFR. However, it is worth pointing out that although the investigation concerns the CEFR levels in relation to subordination, no analysis of the kinds of clauses produced by the informants was reported. Thus, one might wonder what types of subordinate clauses learners on the different levels tend to use. With this in mind, the present study focuses on both linguistic complexity and subordination types.

Prodeau et al. (2012) examined the CEFR scale in relation to developmental stages for L2 French proposed in previous research, and aimed at finding linguistic features in French B1 and B2 texts that were characteristic for the respective CEFR levels. The data

consisted of 40 written productions (20 from texts each level), which were written by learners of L2 French with various mother tongues, and taken from a standard graded test known as the *Test de Connaissance du Français*. The authors identified both deviant and target-like morpho-syntactic features in all the texts, focusing on VP and NP morphology. They then compared the samples systematically as a means to determine both which features were present in the productions from both levels and the characteristics of the B2 texts only. Contrary to Forsberg and Bartning (2010), Prodeau and colleagues did not identify any grammatical features that were typical of only one level of the CEFR. The investigation by Prodeau et al. (2012) offers an interesting hypothesis regarding texts rated according to the CEFR scale, namely that French B1 and B2 samples are relatively homogeneous. It should however be noted that the study only concerns two main CEFR levels (B1 and B2). Although the authors motivate this by arguing that it is at these levels that grammatical features start to emerge in earnest, it is possible that focusing on more than two levels of the framework could yield more fine-grained results. Thus, the present study concerns written output from three main CEFR levels: A2, B1 and B2.

Summing up, we can conclude that Gyllstad et al. (2014) did not look at the types of clauses included in their data. Further, Prodeau et al. (2012) only considered two out of six main CEFR levels, whereas the studies by Rosenberg (2011) and Forsberg and Bartning (2010) concern texts written in one rather than several languages. With this in mind, the current investigation seeks to examine the linguistic accuracy and complexity of texts from three main CEFR levels written in two different languages. More precisely, the following research questions will be addressed based on the hypothesis that linguistic complexity and accuracy increases with higher levels of the CEFR scale:

- What are the differences between the CEFR levels A2, B1 and B2 within the same language regarding subject-verb agreement and subordination, and what do these differences indicate in terms of linguistic complexity and accuracy?
- What are the differences between the same CEFR levels across different languages regarding subject-verb agreement and subordination, and what do these differences indicate in terms of linguistic complexity and accuracy?

3. Methods and materials

3.1 Participants

The data used in this study consists of texts in L2 English (N = 24) and L3 French (N=24) written by students enrolled in different schools in a city in the south of Sweden. The participants were at two different levels of the Swedish school system: year nine and year twelve (i.e. the final year of upper-secondary school). The students were thus between 15 and 19 years of age. Table 2 below provides additional information about the learners whose texts were analysed in this study. As stated earlier, the data were originally collected and used in an investigation by Gyllstad et al. (2014) (see 2.5), and the procedures outlined in the next subsection were carried out in connection to that study.

Table 2: The participants from the two target languages

	Mean age (SD)	Gender (%)		L1
		m	f	
L2 English	16.3 (1.27)	37.5	62.5	Swedish
L3 French	16.6 (1.32)	25	75	Swedish

3.2 The tasks

The written samples analysed in the present study are based on two different tasks. The first assignment was to write an email message to a teacher, and the second task was to write a story about a nice or exciting event. All the texts were written on a computer online, as this was thought to “have more appeal [to the participants] than a traditional paper-and-pencil format” (Gyllstad et al., 2014, p. 11).

Once the data had been collected, the texts were assessed by experienced CEFR raters. Two raters read the English texts whereas two other raters assessed the French samples, and they all did so independently of each other. The raters were asked to read the written samples and give each text a CEFR rating between A2 and C1 using a CEFR scale that they had been given (see Appendix 1). When assessing the texts, the raters were also instructed to indicate the degree of certainty of each score using a scale ranging from *completely certain* to

completely uncertain. If they were not sure of whether they had assessed a text correctly, the raters were asked to provide a second CEFR score as an alternative to the rating that they had given the sample in question.

The data used and collected by Gyllstad and colleagues (2014) consists of texts written by 120 unique informants. Out of these productions, 48 samples were included in the present analysis. The 48 texts were chosen because together, they form a data set with an equal number of samples from each of the targeted CEFR levels across languages (e.g. the same number of English B2 texts and French B2 samples). To the greatest possible extent, I also aimed at analysing texts that had been given the same CEFR score by both raters as a means to ensure that the selected texts corresponded to the CEFR level that they had been placed on. Table 3 below illustrates the distribution of the samples, and indicates that 10 A2 texts, 10 texts on the B1 level and four B2 samples were analysed in each language (i.e. French and English). The choice to only include a total of eight B2 texts is motivated by the fact that out of the 38 French productions from Gyllstad et al.'s (2014) original data set, only four samples were rated B2. Moreover, it was not always possible to analyse an equal number of narratives and e-mails from each CEFR level because, for example, only two of the English A2 texts available were narrative texts. As can be seen in Table 3, this led me to include eight e-mails and two stories from this level in the analysis.

In the current investigation, the texts that were given different CEFR scores by the two raters were classified as belonging to the lower of these levels, as it was assumed that if a sample was given, say, one A2 rating and one rating of B1, it could at least be considered an A2 text. The productions that straddled more than two CEFR levels (e.g. texts that were given a rating of A2 by the first rater and B2 by the second rater) were excluded from the analysis.

Table 3: The data set analysed in the present investigation

	L2 English		L3 French	
	Text 1 (e-mail)	Text 2 (story)	Text 1 (e-mail)	Text 2 (story)
A2	8	2	5	5
B1	5	5	6	4
B2	2	2	3	1

3.3 Coding and analysis of the data

Once the written samples mentioned above had been selected, all the texts were converted into basic CHAT-format by means of the CLAN toolbox (MacWhinney, 2000). This allowed me to code the texts for T-units and clauses as a means to measure the linguistic complexity of the productions and look at the distribution of different clause-types in the data. Following Gyllstad et al. (2014), a main clause and all its potential associated subordinate clauses were seen as one T-unit². An English example of a T-unit is given in (8) and (9) exemplifies T-unit in French. Coordinated sentences like the example in (10) were seen as two different T-units and in run-on sentences such as (11), the comma was replaced by a full stop. The two parts of the run-on sentence were then interpreted as two separate sentences (i.e. two different T-units). All the examples below are taken from the data set analysed in the present investigation.

(8) [T-UNIT [MATRIX I really feel] [SUBCLAUSE that it's the best place on earth]] (EN_B1_SF_text2).

(9) [T-UNIT [MATRIX Je ne crois pas] [SUBCLAUSE que je pourrai faire cette composition]]
I do not think that I will be able to take this written exam
(FR_B2_TA_text1).

(10) [T-UNIT The sun was shining] [T-UNIT and the birds were singing] (EN_B2_KN_text2).

(11) I hope you can answer my questions, I'm waiting for an answer (EN_B1_SF_text1).

In order to measure the accuracy of the written texts and thus partially address the research questions included in this essay (see 2.5), a second analysis focusing on subject-verb agreement was carried out. For practical reasons, the analysis only dealt with main verbs and auxiliary verbs in the present tense. I thus went through all the texts manually and entered the total number of instances requiring subject-verb agreement in the present tense and the number of correctly conjugated verbs among these instances in a spreadsheet, as a means to retrieve the percentage of accurate present tense verb forms in the texts. When I was uncertain

² In my analysis, following Quirk, Greenbaum, Leech and Svartvik (1985), the term *matrix clause* is used to designate the superordinate clause minus its subordinate clause(s). Furthermore, following Gyllstad et al (2014), subordinate clauses are excluded when looking at main clauses, although they are part of the main clause from a grammatical (i.e. clause functional) point of view (Biber et al., 1999; Quirk et al., 1985).

of the correct form of a verb I consulted the *Longman Grammar of Spoken and Written English* or *Bescherelle: La conjugaison pour tous* (see References).

4. Results

4.1 Differences in complexity between the CEFR levels

4.1.1 Complexity measures

The aim of this study is to investigate how texts on different CEFR levels within the same language as well as samples from the same CEFR levels across languages differ in terms of linguistic accuracy and complexity. Following Gyllstad et al. (2014), the complexity of the samples in the analysed data set was determined by looking at the mean number of words per T-unit and clause in the texts. The analysis was based on the group means provided in Table 4 below.

Table 4: The mean length of T-unit and mean length of clause in the data from the targeted CEFR levels

Group	1. Mean length of T-unit		2. Mean length of clause	
	Mean	SD	Mean	SD
ENG A2	9.9	2.4	6.0	.77
ENG B1	10.5	1.4	6.9	.81
ENG B2	9.3	1.0	7.0	.86
FRA A2	8.0	1.8	6.2	1.10
FRA B1	8.6	1.9	6.6	1.16
FRA B2	10.8	1.7	6.8	.96

When looking at Table 4, we can, for example, see that the mean length of T-unit (i.e. mean number of words per T-unit) is 9.9 words in the English A2 texts and 8.0 words in the French samples from the same CEFR level. Table 4 also shows that the mean number of words per clause is 6.9 words in the English texts given a rating of B1, and 6.6 words in the French B1 samples. The mean values reported in Table 4 are illustrated in Figure 1 and 2 below. Figure 1 shows error bars for measure 1 (i.e. mean length of T-unit) and Figure 2 graphically illustrates the second complexity measure (i.e. mean length of clause).

Figure 1: Error bar graphs indicating the mean length of the T-units in the analysed texts

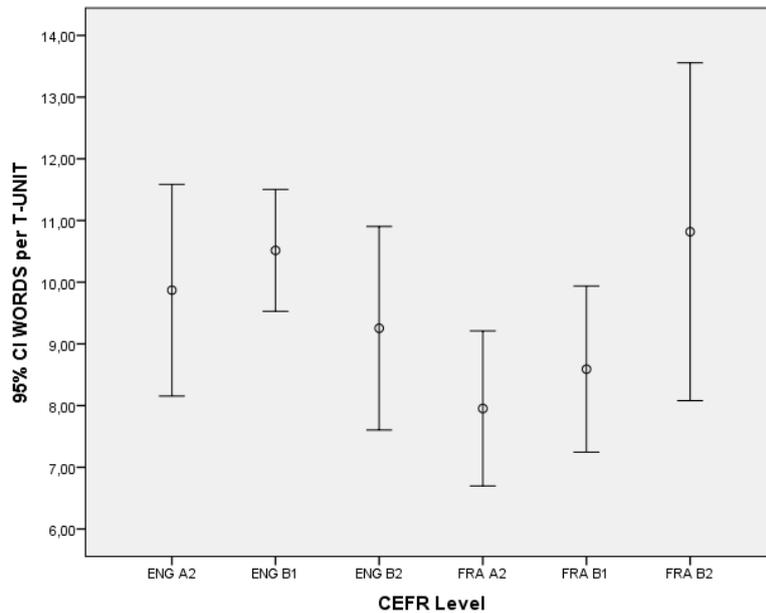
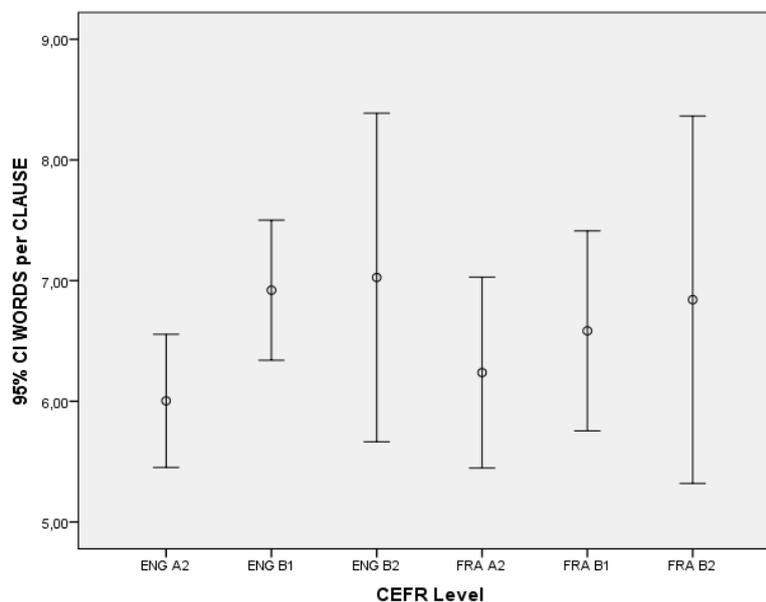


Figure 2: Error bar graphs indicating the mean length of the clauses in the analysed texts



In Figure 1 and 2 above, the circle in the middle of each vertical bar represents the mean of the group, and the bar itself shows the confidence interval around that mean. For example, the sample mean in Figure 1 for the group ENG A2 is 9.9 and the estimated population mean is

likely to fall between about just over 8 and just under 12. This means that in 95 times out of 100 the population mean would be between 8 and 12, hence the wording on the Y-axis that says “95% CI” (CI stands for *confidence interval*). For most error bars in Figure 1 and Figure 2, there is considerable overlap, which means that in many cases, the differences in complexity between the respective CEFR categories are not expected to be statistically significant. In order to verify the results illustrated in Figure 1 and 2, a statistical test referred to as an ANOVA was carried out. The results are shown in Table 5 below:

Table 5: Overview of the ANOVA

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
WORDS per text	Between Groups	60796,867	5	12159,373	14,387	,000
	Within Groups	35497,050	42	845,168		
	Total	96293,917	47			
WORDS per T-UNIT	Between Groups	50,205	5	10,041	2,996	,021
	Within Groups	140,762	42	3,351		
	Total	190,967	47			
WORDS per CLAUSE	Between Groups	6,543	5	1,309	1,399	,244
	Within Groups	39,288	42	,935		
	Total	45,830	47			

The results presented in Table 5 indicate that the words per T-unit measure yielded significant differences as the Sig. column has a value that is less than .05 (.021). The second complexity measure (words per clause), on the other hand, was statistically insignificant as the Sig. value is .244, and thus well above the conventional cut-off of .05. In order to find out where the statistical differences were (i.e. between which CEFR categories) a post-hoc test was used. It revealed that the group means from ENG B1 and FRA A2 for mean length of T-unit were significantly different, $F(5, 42) = 2.99, p < .05$. For the mean words per clause measure, no statistically significant differences between means were observed when checking with the post-hoc test, and we did not expect that since the values in Table 5 did not indicate this anyway.

4.1.2 Analysis of subordination types

In section 2.5, we saw that Gyllstad et al. (2014) looked at the subordinate clauses produced by students on different CEFR levels without presenting an analysis of the different subordination types in their data, even though clause-types constitute an important aspect of subordination as a grammatical phenomenon (Biber et al., 1999). With this gap in mind, the present investigation aims at determining what kinds of subordinate clauses learners on the targeted CEFR levels within the same language and across different languages typically use. The following table (6) therefore indicates the frequency of the clause-types exemplified and summarized in 2.4.1 in the analysed productions. The purpose of the subordination type analysis presented below is not to measure the complexity of the current data set. Instead, the aim is to present a more qualitative analysis of the written texts, and it is assumed that this type of qualitative differences between CEFR levels and languages are important to map out in the empirical underpinning of the CEFR system.

Table 6: The frequency of the different types of subordinate clauses produced by the informants

CEFR level	Adverbial clauses		Nominal clauses		Relative clauses		Other clauses	
	n	%	n	%	n	%	n	%
English A2	20	43.5	19	41.3	7	15.2	0	0
English B1	33	39.8	31	37.3	18	21.7	1	1.2
English B2	8	22.2	24	66.7	4	11.1	0	0
French A2	19	63.3	9	30.0	2	6.7	0	0
French B1	17	37.0	23	50.0	6	13.0	0	0
French B2	7	26.9	12	46.2	7	26.9	0	0

Table 6 illustrates that out of all the subordinate clauses in the French A2 texts, 63.3 % were adverbial clauses such as the temporal clause in (12), the clause expressing a cause in (13) and the conditional clause in (14). In the French samples rated B1 or B2, on the other hand,

nominal clauses was the most common subordination type, as they constitute 50 % of the subordinate clauses from the B1 texts and 46.2 % of the B2 clauses. Examples of nominal clauses from the French data are given in (15) and (16) below. More precisely, the example in (15) is a complement clause and (16) is an interrogative clause. Table 6 also shows that nominal clauses like the examples in (17) and (18) constitute 66.7 % of the dependent clauses found in the English samples given a rating of B2. Example (17) is a complement clause and (18) is an interrogative clause. In the examples (12)- (18), the targeted features have been underlined, and no attention has been paid to possible spelling mistakes, morpho-syntactic deviances etc.

(12) Quand j'étais petite, j'avais une petite bateau (FR_A2_AK_text2).

When I was little, I had a small boat.

(13) Je n'ai pu pas lire les pages au livre de text parce que vous n'avez pas pu me donner le livre (FR_A2_AW1_text 1).

I have not been able to read the pages in the textbook, because you have not been able to give me the book.

(14) Si le test est demain, peut-je le faire une autre jour (FR_A2_ES1_text1)?

If the test is tomorrow, can I take it another day?

(15) Je sais que nous avons un teste (FR_B1_OL_text1).

I know that we have a test.

(16) [M]ais je ne sais pas exactement de quoi il s'agit (FR_B2_JL_text1).

[B]ut I do not know exactly what it is about.

(17) I have heard that the other classes have started reading a novel (EN_B2_AN2_text1).

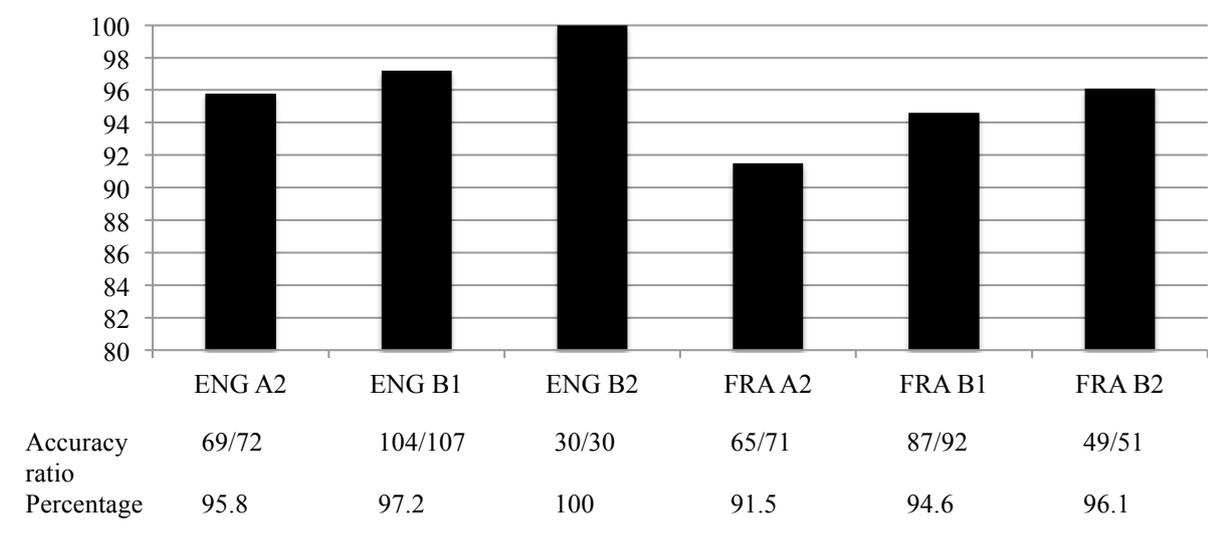
(18) Also, I wonder what the next task will be, after we have finished this one (EN_B2_AN2_text1).

4.2 Differences in accuracy between the CEFR levels

In order to address the parts of the research questions in the essay concerning linguistic accuracy, the proportion of correctly conjugated present tense verbs in the data was determined by dividing the number of correct (present tense) verb forms with total number of instances requiring subject-verb agreement in the present tense. The top half of Figure 3 below graphically illustrates the percentage of accurately inflected verbs in all the texts

combined from the respective CEFR levels, and the bottom half shows the accuracy ratio and the proportion of correct present tense verb forms in the texts.

Figure 3: The proportion of correct present tense verb forms in the texts



According to Figure 3, there is a general increase in the percentage of correctly conjugated verbs produced by the groups. For example, the proportion of correct (present tense) verb forms is 95.8 % in all the English A2 texts combined and 97.2 % in the texts in English rated B1. When looking at the same CEFR levels across different languages we can see that 100 % of the verbs in the English B2 texts and 96.1 % of the verbs in the French texts on the same level were inflected correctly. However, when interpreting Figure 3 one must remember that there are more verb forms to choose from in French than in English (see sections 5 and 2.4.2), and that the vertical axis in the first half of the graph starts at 80 as a means to clearly illustrate the findings. In the next section, the present results will be discussed in light of the information provided in 2.1-2.4 and the previous studies summarized in 2.5.

5. Discussion

This essay addresses the relationship between linguistic accuracy and complexity and two different categories of written output: texts from different CEFR levels within the same

language and samples from the same CEFR levels across different languages. Starting with accuracy, the present findings suggest that it generally increases with higher CEFR levels. For example, we saw in section 4.2 that 97.2 % of the present tense verb forms in the English B1 texts were correct, whereas the B2 texts had a success rate of 100 %. Not surprisingly, this observation is in line with the results obtained by Forsberg and Bartning (2010). As to the differences between the same CEFR levels across the targeted languages, the English texts were generally more accurate than the French samples on the corresponding levels. For instance, the success rate was 95.8 % in the English A2 samples and 91.5 % in the French productions from the same CEFR level. This was expected, because as mentioned earlier, VP morphology is a more complex grammatical phenomenon in French than in English.

The results also indicate that all the participants combined generally mastered subject-verb agreement in the present tense well, as the success rates were typically very high. (see Figure 3). This is in accordance with Rosenberg's (2011) findings as she, among other things, observed that the fifteen-year-old learners in her study inflected 95 % of the verbs in the present tense correctly. It is however important to take into account that the number of different verbs used in both the French and English samples from the current investigation was somewhat limited. For example, the French learners on all levels often produced frequent verbs like *avoir* (*to have*) as exemplified in (19) below, whereas lexical verbs were less common, which corroborates Rosenberg's (2011) findings. Further, when looking at the French lexical verbs that were found in the present data set, there is a difference in that the accurate lexical verb forms used in the French texts rated A2 were typically regular verbs ending in *-er*, whereas the B2 learners generally also conjugated irregular verbs correctly. Compare, for instance, (20) and (21) below. In example (20), which is from an A2 text, the writer provides the correct form of the regular verb *espérer* (*to hope*) but uses an erroneous form of *pouvoir* (*to be able to*), which is irregular, whereas (21) was found in a sample rated B2 and illustrates accurate usage of the irregular verb *croire* (*to think*). This seems to suggest that learners use irregular verbs more accurately as they become more proficient, which is in line with both Rosenberg's (2011) and Bartning and Schlyter's (2004) observations.

(19) J'ai deux questions pour vous [...] (FR_A2_CB_text1)

I have two questions for you [...]

(20) *J'espère que je peut retourner [...] (FR_A2_CB_text1)

I hope that I can come back [...]

(21) Je ne crois pas que je pourrai faire cette composition [...] (FR_B2_TA_text1)

I do not think that I will be able to take this written exam [...]

The learners of English were also generally successful when conjugating verbs in the present tense. The difficulty to employ the third person singular –s noted by Källkvist and Petersson (2006) (see 2.4.2) was not observed in the present study, which raises the question of how the high success rates can be explained. One possibility is that the participants used constructions that they were familiar with in order to be as accurate as possible. Such a tendency was observed by Hasselgren (1994), who talks about so-called *lexical teddy bears*, i.e. familiar words that learners cling to when expressing themselves. Thus, it is possible that the accuracy observed in the current investigation is due to the fact that the nature of the task permitted the informants to employ “*grammatical teddy bears*” in the form of verbs where they were not forced to pay attention to the employment of the third person -s.

Returning to lexical complexity, the present findings show that out of the six CEFR categories that were analysed in this study (i.e. English A2, English B1, English B2, French A2, French B1 and French B2), only two groups were significantly different in regard to mean length of T-unit: French A2 and English B1. Thus, the results are in line with the observations made by Prodeau et al. (2012) in that they indicate a difficulty in differentiating between the CEFR levels. It should however be taken into account that Prodeau and colleagues systematically compared qualitative features in B1 and B2 texts, without focusing on the complexity measures used in the present study. Regardless, the lack of discriminatory features in the current investigation is surprising, as it does not corroborate the results obtained by Gyllstad and colleagues (2014), which indicated significant differences in complexity between the CEFR levels across all the targeted languages (see 2.5). However, it should be stressed that when measuring linguistic complexity, Gyllstad et al. (2014) categorized their data according to the three theoretical CEFR levels A B and C. In contrast to the present analysis, the investigation by Gyllstad et al. (2014) included both A1 and A2 texts, which were merged into category A, whereas the B1 and B2 samples constituted category B etc. Thus, the contradictory results may be explained by the fact that the data were classified differently in the two studies, since it is plausible that more differences emerge when comparing A1 and A2 samples combined to B1 and B2 texts than when comparing, for instance, A2 samples and B1 texts only.

It is also noteworthy that although the present investigation did not reveal any statistically significant differences between the mean number of words per T-unit in the

English B1 and B2 samples, the mean value for length of T-unit was higher in the first of these two categories from a purely descriptive point of view (see 4.1.1). This observation is not in accordance with the results reported by Gyllstad et al. (2014) and does not corroborate the hypothesis that linguistic complexity increases with higher CEFR levels. However, it is important to remember that only four out of the 17 English B2 texts in Gyllstad et al.'s (2014) original data set were analysed in the present study (see 3.2). Out of these four samples, two texts were email messages, which might have affected the outcome of the analysis, as one might not expect learners (regardless of their proficiency level) to use a lot of subordinated sentences in such a text. It is thus possible that the unexpected (but statistically insignificant) decrease in complexity between the levels B1 and B2 in English observed in this study is related to both the number and types of written productions included in the analysis.

Moreover, we saw in 2.2 that recent CAF-related studies focusing on how complexity, accuracy and fluency interact indicate that learners sometimes are unable to produce output that is as accurate as it is complex. Thus, seeing that the English B2 learners conjugated all the verbs in the present tense correctly (see 4.2), another reason why these texts were not significantly more complex than the English samples from the other CEFR levels might be that the observed accuracy occurred at the cost of linguistic complexity. The present investigation thus supports the hypothesis that the three components of the CAF triad are closely connected.

As to the different types of clauses produced by the learners, 41.3 % of the dependent clauses in the English A2 samples from this study were of the nominal type, which is not surprising since Hawkins and Filipović (2012) suggest that finite complement clauses are attested at this level. However, no examples of *ing*-clauses preceding the main clause were found in the present data set, although Hawkins and Filipović (2012) consider them to be one of the characteristics of English B2 texts. The absence of this clause-type might be due to the fact that only four English productions given a rating of B2 were included in the analysis. Further, 63.3 % of the clauses in the French A2 texts were adverbial clauses. This was expected, seeing that Bartning and Schlyter (2004) propose that adverbial clauses are among the first to emerge as a learner of French as a second language acquires subordination. Bartning and Schlyter (2004) also suggest that usage of the *gérondif* and clauses introduced by *donc* (*so*) characterises advanced L2 French (see 2.4.1). This was corroborated in the present investigation, as the only observed instances of these features were found in French B2 samples, which in turn suggests that the *gérondif* illustrated in (22) and the employment of

donc exemplified in (23) are typical B2 features. In the examples (22)-(24) below, all possible grammatical and lexical deviances have been disregarded.

(22) En sortant du train à la gare [...], j'ai vu mon amie (FR_B2_JL_text2).

When I went out of the train at the train station [...], I saw my friend.

(23) Moi j'ai [...] de chose à faire dans tous les cours maintenant, donc je suis un peu stressé

I have [...] things to do in all the courses right now, so I am a little bit stressed

(FR_B2_TA_text1).

(24) What pages shall we read to the test? And shall we do something else then just write

(EN_A2_EL_text1)?

Finally, this essay is limited in several ways, which I will take into account in the future. For example, we saw earlier that the number of texts from each CEFR level included in the analysed data set was relatively small, which might have affected the results. In addition, several of the selected samples were email messages in which the students were explicitly instructed to include questions (see Appendix 2). This typically resulted in consecutive direct questions as exemplified in (24) above, which possibly influenced the complexity of the written productions. Furthermore, there was no second rater of the data, which may have had an impact on the results. With these limitations in mind, it would be interesting to replicate this study using a larger data set containing text-types where subordination can be expected (e.g. narratives only), and thus further contribute to the much-needed body of empirical CEFR-related studies.

6. Conclusion

The study reported in this essay sought to examine texts written by Swedish learners of L2 English and L3 French, which had been assessed according to the *Common European Framework of Reference for Languages* (CEFR). The aim was to determine the differences between output from (a) the respective CEFR levels within the same language and (b) the same CEFR levels across different languages, with regard to subject-verb agreement and

subordination. It was hypothesized that the usage of subordination and the proportion of correctly conjugated verbs in a text are indicators of linguistic complexity and accuracy respectively, and that the accuracy and complexity demonstrated in the data would increase with higher CEFR levels.

The research questions were addressed by means of written samples in English (N=24) and French (N=24) in the form of email messages and stories. Following Gyllstad et al. (2014), the complexity of the texts was determined by looking at the mean length of T-unit and the mean number of words per clause in the data. Two additional analyses concerning the types of clauses produced by the learners and the proportion of correct present tense verb forms in the samples were also carried out.

In sum, no statistically significant differences were found between the targeted CEFR categories when measuring the mean length of the clauses in the texts. Similarly, only the French A2 samples and the English B1 texts were significantly different in terms of mean length of T-unit, which indicates that all the texts were rather homogenous with regard to linguistic complexity. These findings are not in line with previous observations and may, among other things, be explained by the relatively small number of analysed productions. However, the results also showed that the higher the CEFR rating a text had been given, the more accurate it was. Finally, the clause-type analysis revealed that, for example, a majority of the subordinate clauses found in the French A2 samples were of the adverbial type, which is in accordance with previous SLA research.

References

- Bartning, I., & Schlyter, S. (2004) Itinéraires acquisitionnels et stades de développement en français L2. *French language studies*, 14, 281-299.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (Eds.). (1999). *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education Limited.
- Brossier, E. (Ed.). (2006). *Bescherelle: La conjugaison pour tous*. Paris: Hatier.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Forsberg, F., & Bartning, I. (2010). Can linguistic features discriminate between the communicative CEFR-levels? A pilot study of written L2 French. In I. Bartning, M. Martin, & I. Vedder (Eds.), *Communicative proficiency and linguistic development: Intersections between SLA and language testing research* (pp. 133-157). Amsterdam: John Benjamins.
- Gunnarsson, C. (2012). The development of complexity, accuracy and fluency in the written production of L2 French. In A. Housen, F. Kuiken & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 247-276). Amsterdam: John Benjamins.
- Gyllstad, H., Granfeldt, J., Bernardini, P., & Källkvist, M. (2014). Linguistic correlates to communicative proficiency levels of the CEFR: The case of syntactic complexity in written L2 English, L3 French and L4 Italian. In L. Roberts, I. Vedder & J.H. Hulstijn (Eds.), *EUROSLA Yearbook 14* (pp. 1-30). Amsterdam: John Benjamins.
- Hasselgren, A. (1994). Lexical teddy bears and advanced learners: A study into the ways Norwegian students cope with English vocabulary. *International Journal of Applied Linguistics*, 4(2), 237-260.
- Hawkins, J.A., & Filipović, L. (2012). *Criterion features in L2 English: Specifying the Reference Levels of the Common European Framework*. Cambridge: Cambridge University Press.
- Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30(4), 461-473.
- Housen, A., Kuiken, F. & Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurement and research. In A. Housen, F. Kuiken & I. Vedder. (Eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA* (pp. 1-20). Amsterdam: John Benjamins.
- Hulstijn, J.H. (2007). The shaky ground beneath the CEFR : Quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal*, 91, 633-667.
- Hulstijn, J.H., Schoonen, R., de Jong, N.H., Steinel, M.P., & Florijn, A. (2011). Linguistic competences of learners of Dutch as a second language at the B1 and B2 levels of speaking proficiency of the Common European Framework of Reference for Languages (CEFR).

- Language Testing*, 29(2), 203-221.
- Kuiken, F., & Vedder, I. (2012). Syntactic complexity, lexical variation and accuracy as a function of task complexity and proficiency level in L2 writing and speaking. In A. Housen, F. Kuiken & I. Vedder. (Eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA* (pp. 143-170). Amsterdam: John Benjamins.
- Källkvist, M., & Petersson, S. (2006). An *s*, or not an *s*; that is the question: Swedish teenage learners' explicit knowledge of subject-verb agreement in English. In J. Einarsson, E. Larsson Ringqvist & M. Lindgren (Eds.), *Approaches to teaching and learning in linguistic research: Papers from the ASLA symposium in Växjö, 10-11 November 2005* (pp. 112-140). Uppsala: ASLA.
- Little, D. (2007). The common European framework of reference for languages: Perspectives on the making of supranational language education policy. *The Modern Language Journal*, 91, 645-655.
- Martin, M., Mustonen, S., Reiman, N., & Seilonen, M. (2010). On becoming an independent user. In I. Bartning, M. Martin & I. Vedder (Eds.), *Communicative proficiency and linguistic development: Intersections between SLA and language testing research* (pp. 57-81). Amsterdam: John Benjamins.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah: Lawrence Erlbaum Associates.
- Norris, J.M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30(4), 555-578.
- Ortega, L. (2009). *Understanding second language acquisition*. London: Hodder Education.
- Prodeau, M., Lopez, S., & Véronique, D. (2012). Acquisition of French as a second language: Do developmental stages correlate with CEFR levels?. *Apples- Journal of Applied Language Studies*, 6(1), 47-68.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London: Longman
- Riegel, M., Pellat, J.C., & Rioul, R. (Eds.). (1994). *Grammaire méthodique du français*. Paris: Presses Universitaires de France.
- Rosenberg, J. (2011). *Existe-t-il des corrélations entre les stades de développement morphosyntaxique et les niveaux proposés du Cadre européen commun de référence ? : Une étude de productions écrites par des apprenants suédophones de français L2* (Unpublished Bachelor thesis). Lund University, Sweden.
- Selinker, L. (1972). *Interlanguage*. *IRAL*, 10, 209-231.
- Tarone, E. (2006). *Interlanguage*. In K. Brown (Ed.), *Encyclopaedia of language and linguistics* (pp. 747-752). Boston: Elsevier.
- Wall, K., Ekman, M., Béhar, D., & Kroning, H. (Eds.). (1999). *Bonniers Franska Grammatik*.

Stockholm: Bonnier Utbildning.

Ågren, M. (2008). *À la recherche de la morphologie silencieuse: Sur le développement du pluriel en français L2 écrit* (Doctoral dissertation). Retrieved from SwePub. (1267092)

Ågren, M., Granfeldt, J., & Schlyter, S. (2012). The growth of complexity and accuracy in L2 French: Past observations and recent applications of developmental stages. In A. Housen, F. Kuiken & I. Vedder. (Eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA* (pp. 95-120). Amsterdam: John Benjamins

Appendix 1: The rating scale used by the raters, compiled from several CEFR scales

	OVERALL WRITTEN PRODUCTI ON	WRITTEN INTERACTION	CORRESPONDENCE & NOTES, MESSAGES, FORMS	CREATIVE WRITING & THEMATIC DEVELOPMENT & COHERENCE AND COHESION
A1	Can write simple isolated phrases and sentences.	Can ask for or pass on personal details in written form.	Can write a short simple postcard. Can write numbers and dates, own name, nationality, address, age, date of birth or arrival in the country, etc. such as on a hotel registration form.	Can write simple phrases and sentences about themselves and imaginary people, where they live and what they do. Can link words or groups of words with very basic linear connectors like 'and' or 'then'.
A2	Can write a series of simple phrases and sentences linked with simple connectors like 'and', 'but' and 'because'.	Can write short, simple formulaic notes relating to matters in areas of immediate need.	Can write very simple personal letters expressing thanks and apology. Can take a short, simple message provided he/she can ask for repetition and reformulation. Can write short, simple notes and messages relating to matters in areas of immediate need.	Can write about everyday aspects of his/her environment, e.g. people, places, a job or study experience in linked sentences. Can write very short, basic descriptions of events, past activities and personal experiences. Can write a series of simple phrases and sentences about their family, living conditions, educational background, present or most recent job. Can write short, simple imaginary biographies and simple poems about people. Can tell a narrative or describe something in a simple list of points. Can use the most frequently occurring connectors to link simple sentences in order to tell a narrative or describe something as a simple list of points. Can link groups of words with simple connectors like 'and', 'but' and 'because'.
B1	Can write straightforward connected texts on a range of familiar subjects within his field of interest, by linking a series of shorter discrete elements into a linear sequence.	Can convey information and ideas on abstract as well as concrete topics, check information and ask about or explain problems with reasonable precision. Can write personal letters and notes asking for or conveying simple information of immediate relevance, getting across the point he/she feels to be important.	Can write personal letters giving news and expressing thoughts about abstract or cultural topics such as music, films. Can write personal letters describing experiences, feelings and events in some detail. Can write notes conveying simple information of immediate relevance to friends, service people, teachers and others who feature in his/her everyday life, getting across comprehensibly the points he/she feels are important. Can take messages communicating enquiries, explaining problems.	Can write straightforward, detailed descriptions on a range of familiar subjects within his/her field of interest. Can write accounts of experiences, describing feelings and reactions in simple connected text. Can write a description of an event, a recent trip – real or imagined. Can narrate a narrative. Can reasonably fluently relate a straightforward narrative or description as a linear sequence of points. Can link a series of shorter, discrete simple elements into a connected, linear sequence of points.
B2	Can write clear, detailed texts on a variety of subjects related to his/her field of	Can express news and views effectively in writing, and relate to those of others.	Can write letters conveying degrees of emotion and highlighting the personal significance of events and experiences and commenting on the correspondent's news and	Can write clear, detailed descriptions of real or imaginary events and experiences, marking the relationship between ideas in clear connected text, and following established conventions of the genre concerned. Can write clear, detailed descriptions on a

	interest, synthesising and evaluating information and arguments from a number of sources.		views.	<p>variety of subjects related to his/her field of interest.</p> <p>Can write a review of a film, book or play.</p> <p>Can develop a clear description or narrative, expanding and supporting his/her main points with relevant supporting detail and examples.</p> <p>Can use a limited number of cohesive devices to link his/her utterances into clear, coherent discourse, though there may be some 'jumpiness' in a long contribution.</p> <p>Can use a variety of linking words efficiently to mark clearly the relationships between ideas.</p>
C1	Can write clear, well-structured texts of complex subjects, underlining the relevant salient issues, expanding and supporting points of view at some length with subsidiary points, reasons and relevant examples, and rounding off with an appropriate conclusion.	Can express him/herself with clarity and precision, relating to the addressee flexibly and effectively.	Can express him/herself with clarity and precision in personal correspondence, using language flexibly and effectively, including emotional, allusive and joking usage.	<p>Can write clear, detailed, well-structured and developed descriptions and imaginative texts in an assured, personal, natural style appropriate to the reader in mind.</p> <p>Can give elaborate descriptions and narratives, integrating sub-themes, developing particular points and rounding off with an appropriate conclusion.</p> <p>Can produce clear, smoothly flowing, well-structured speech, showing controlled use of organisational patterns, connectors and cohesive devices.</p>
C2	Can write clear, smoothly flowing, complex texts in an appropriate and effective style and a logical structure which helps the reader to find significant points.	As C1	As C1	<p>Can write clear, smoothly flowing, and fully engrossing stories and descriptions of experience in a style appropriate to the genre adopted.</p> <p>Can create coherent and cohesive text making full and appropriate use of a variety of organisational patterns and a wide range of cohesive devices.</p>

Appendix 2: The instructions given to the informants

Task 1

Instructions for students in year 9 (translated from Swedish, instructions were identical for learners of English and French): “Email message to your teacher: You have been absent from school for one week. Soon you will sit an English (French) test. Please ask your teacher two questions regarding the test. Ask two questions about what happened over the week when you were away. Remember to begin and to end your email message in an appropriate way. Write in English (French).”

Instructions for students in their final year of upper-secondary school (translated from Swedish, instructions were identical for learners of English and French): “Email message to your teacher: You have been absent from school for one week. You have been given an assignment that is due next week. Send an email to your teacher. Tell your teacher why you have been absent. Ask two questions about the assignment. Ask two questions about the course content. Write in English (French).

Task 2

Instructions for students of English and French in year 9: “Please write a story about something fun, nice or exciting that has happened to you. What happened? Why was the event fun, nice or exciting? Write in English (French).”

Instructions for students of English and French in the final year of upper-secondary school: “Please write about a special event that occurred or a special experience that you have had that you can still remember. What happened? Why is the event or experience memorable? Write in English (French).”

